Convex Functions

Wing-Kin (Ken) Ma
The Chinese University of Hong Kong (CUHK)

Course on Convex Optimization for Wireless Comm. and Signal Proc.
Jointly taught by Daniel P. Palomar and Wing-Kin (Ken) Ma
National Chiao Tung Univ., Hsinchu, Taiwan

December 19-21, 2008

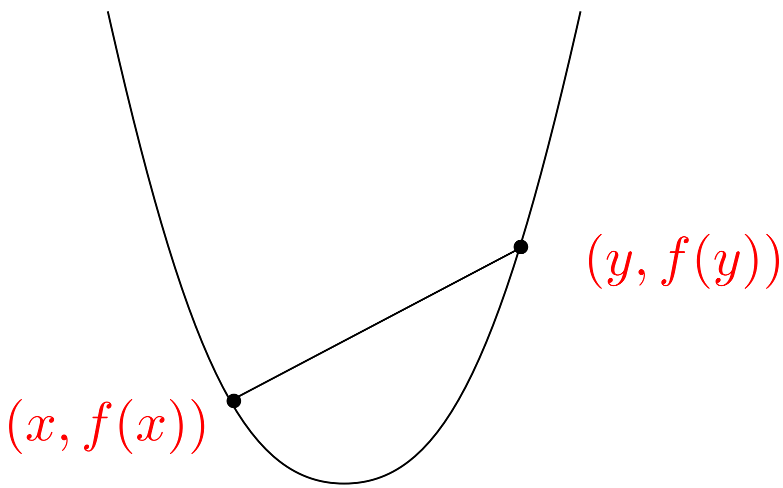
Convex Functions

A function $f : \mathbf{R}^n \to \mathbf{R}$ is said to be **convex** if

i) $\mathbf{dom} f$ is convex; and

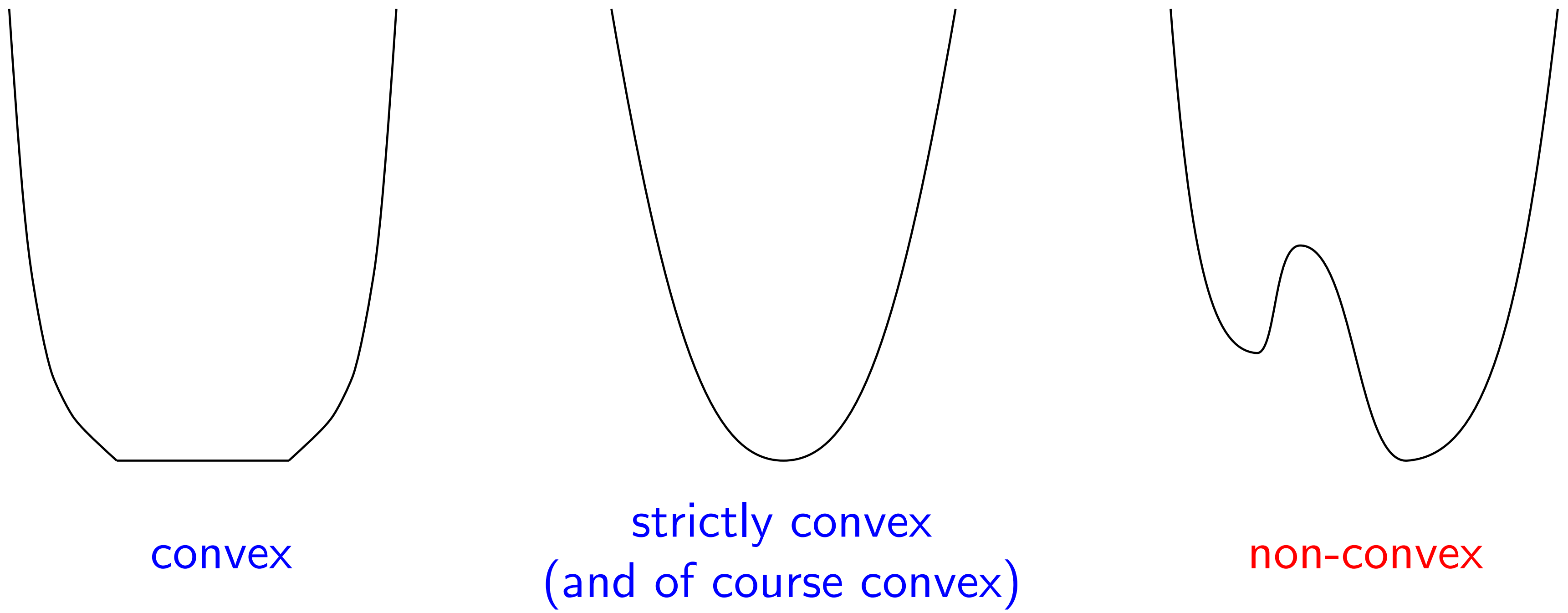
ii) for any $x, y \in \mathbf{dom} f$ and $\theta \in [0, 1]$,

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$



- f is **strictly convex** if $f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$ for all $0 < \theta < 1$ and for all $x \neq y$.

- f is **concave** if $-f$ is convex.



1st and 2nd Order Conditions

- **Gradient** (for differentiable f)

$$\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n}\right]^T \in \mathbf{R}^n$$

- **Hessian** (for twice differentiable f): A matrix function $\nabla^2 f(x) \in \mathbf{S}^n$ in which

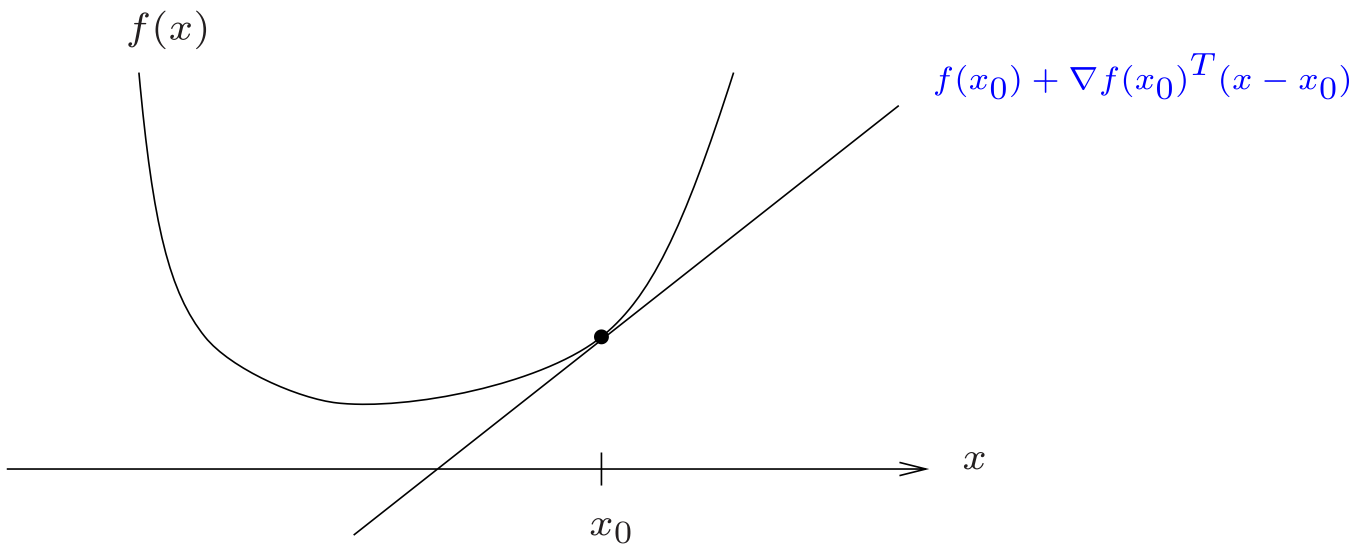
$$[\nabla^2 f(x)]_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

- Taylor series:

$$f(x+\nu) = f(x) + \nabla f(x)^T \nu + \tfrac{1}{2}\nu^T \nabla^2 f(x) \nu + \ldots$$

- **First-order condition:** A differentiable f is convex iff given any $x_0 \in \textbf{dom} f$,

$$f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0), \quad \forall x \in \textbf{dom} f$$



- **Second-order condition:** A twice differentiable f is convex if and only if

$$\nabla^2 f(x) \succeq 0, \quad \forall x \in \textbf{dom} f$$

Examples on $\mathbb{R}$

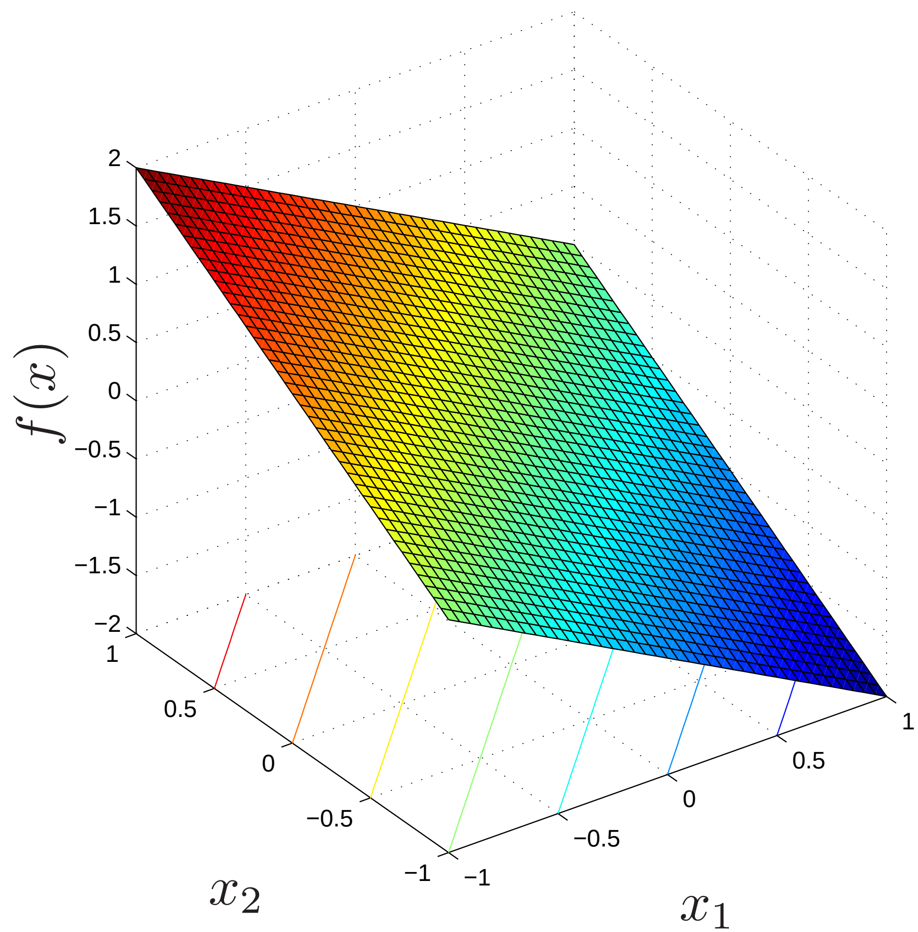
- $ax + b$ is convex. It is also concave.
- x^2 is convex (on $\mathbf{R}$).
- $|x|$ is convex.
- $e^{\alpha x}$ is convex.
- $\log x$ is concave on $\mathbf{R}_{++}$.
- $x \log x$ is convex on $\mathbf{R}_+$.
- $\log \int_{-\infty}^{x} e^{-t^2/2} dt$ is concave.

Examples on $\mathbf{R}^n$

Affine function

$$f(x) = a^T x + b = \sum_{i=1}^{n} a_i x_i + b$$

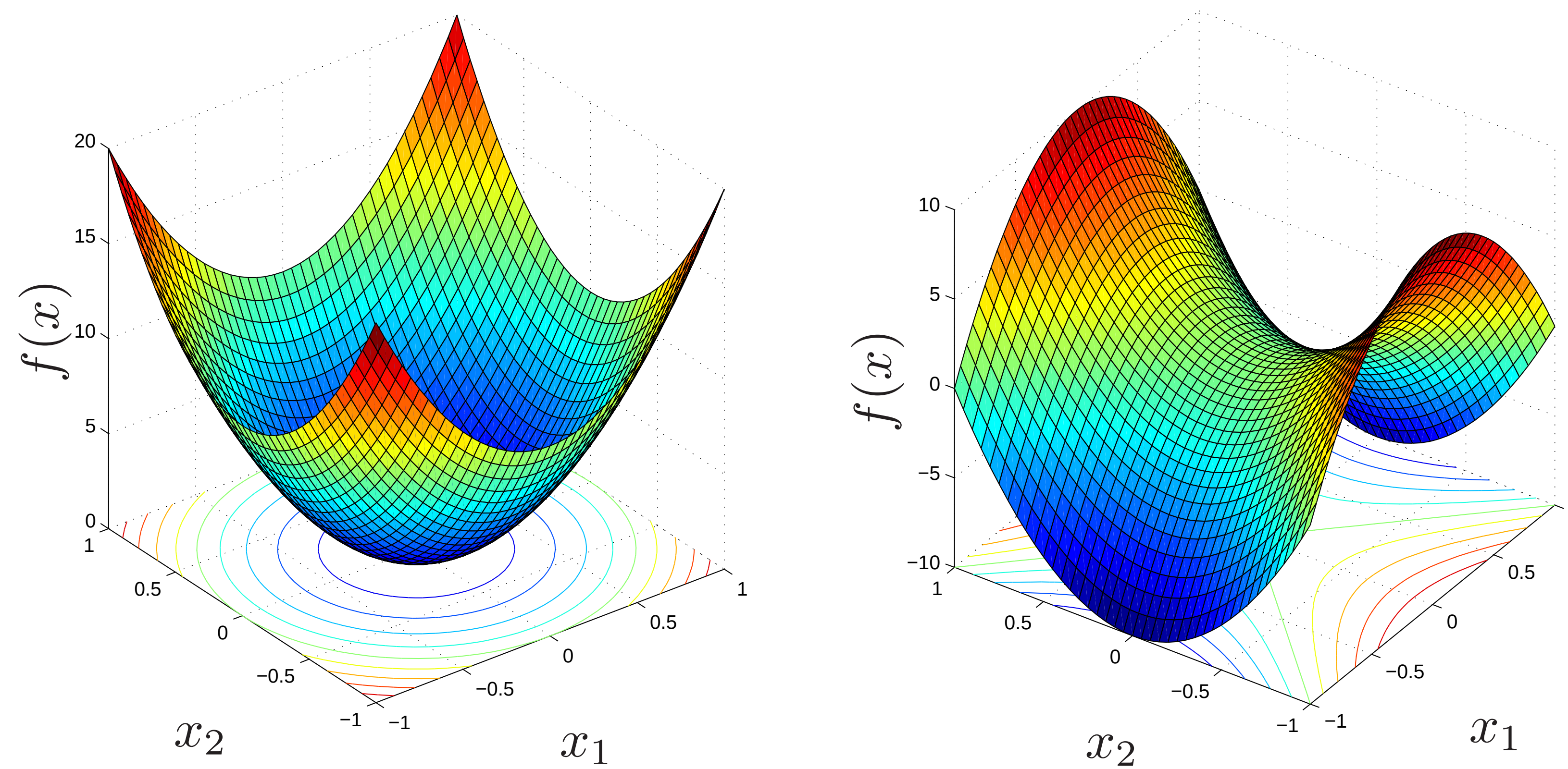
is both convex and concave.



Quadratic function

$$f(x) = x^T P x + 2q^T x + r = \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} x_i x_j + 2 \sum_{i=1}^{n} q_i x_i + r$$

is convex if and only if $P \succeq 0$.

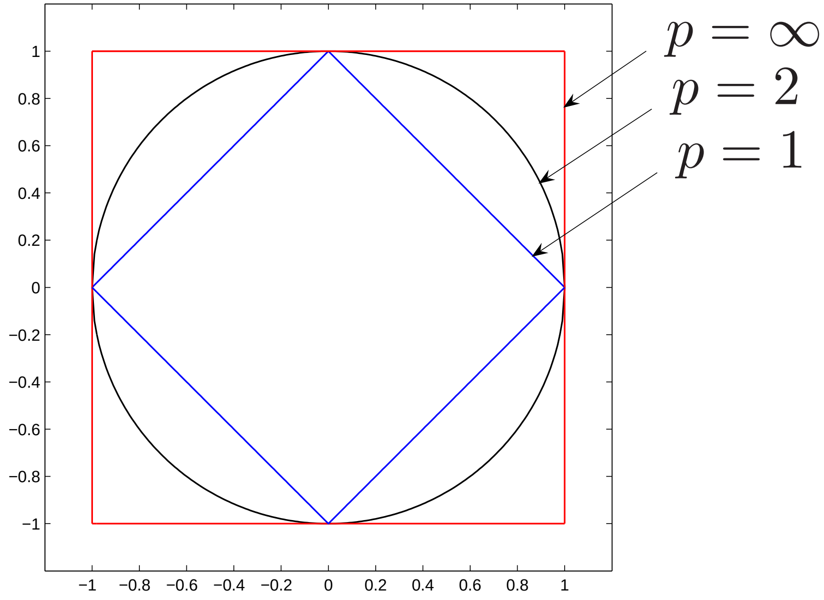


(a) $P \succeq 0$. **(b)** $P \not\succeq 0$.

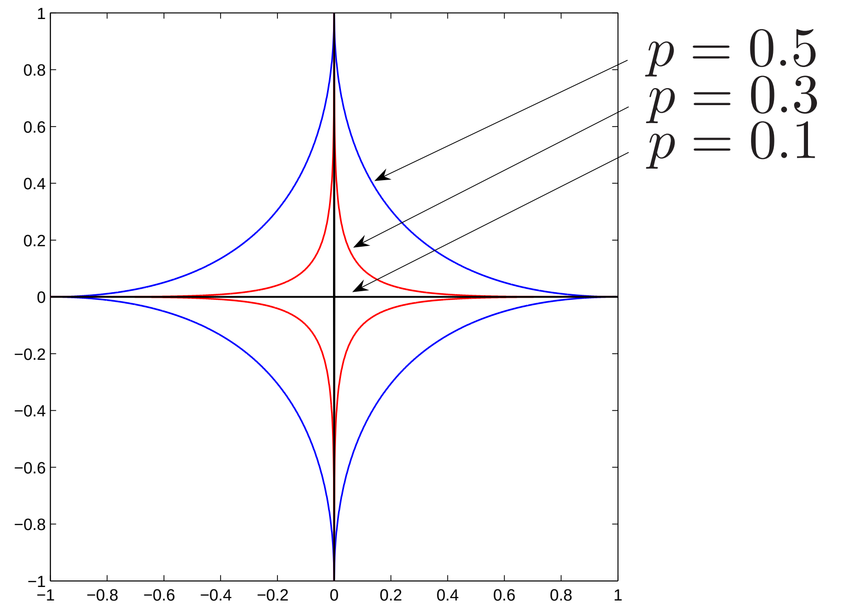
p-norm

$$f(x) = \|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

is convex for $p \geq 1$.



(a) Region of $\|x\|_p = 1$, $p \geq 1$.



(b) Region of $\|x\|_p = 1$, $p \leq 1$.

Geometric mean

$$f(x) = \prod_{i=1}^{n} x_i$$

is concave on $\mathbf{R}^n_{++}$.

Log-sum-exp.

$$f(x) = \log\left(\sum_{i=1}^{n} e^{x_i}\right)$$

is convex on $\mathbf{R}^n$. (Log-sum-exp. can be used as an approx. to $\max_{i=1,\ldots,n} x_i$)

Examples on $\mathbf{S}^n$ (intuitively less obvious)

Affine function:

$$f(X) = \mathbf{tr}(AX) + b = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ji}X_{ij} + b$$

is convex and concave.

Logarithmic determinant function:

$$f(X) = \log\det(X)$$

is concave on $\mathbf{S}^n_{++}$.

Maximum eigenvalue function:

$$f(X) = \lambda_{\max}(X) = \sup_{y \neq 0} \frac{y^T X y}{y^T y}$$

is convex on $\mathbf{S}^n$.

Convexity Preserving Operations

Affine transformation of the domain:

$$f \text{ convex} \Longrightarrow f(Ax+b) \text{ convex}$$

Example: (Least squares function)

$$f(x) = \|y - Ax\|_2$$

is convex, since $f(x) = \|\cdot\|_2$ is convex.

Example: (MIMO capacity)

$$f(X) = \log\det(HXH^T + I)$$

is concave on $\mathbf{S}^n_+$, since $f(X) = \log\det(X)$ is concave on $\mathbf{S}^n_{++}$.

Composition: Let $g : \mathbf{R}^n \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$.

g convex, h convex, extended h nondecreasing $\Longrightarrow f(x) = h(g(x))$ convex

g concave, h convex, extended h nonincreasing $\Longrightarrow f(x) = h(g(x))$ convex

Example:

$$f(x) = \|y - Ax\|_2^2$$

is convex by composition, where $g(x) = \|y - Ax\|_2$, $h(x) = \max\{0, x^2\}$.

Non-negative weighted sum:

$$\begin{array}{l} f_1, \ldots, f_m \text{ convex} \\ w_1, \ldots, w_m \geq 0 \end{array} \implies \sum_{i=1}^{m} w_i f_i \text{ convex}$$

Example: Regularized least squares function

$$f(x) = \|y - Ax\|_2^2 + \gamma \|x\|_2^2$$

is convex for $\gamma \geq 0$.

Extension of non-negative weighted sum:

$$\begin{array}{l} f(x,y) \text{ convex in } x \text{ for each } y \in \mathcal{A} \\ w(y) \geq 0 \text{ for each } y \in \mathcal{A} \end{array} \implies \int_{\mathcal{A}} w(y) f(x,y) dy \text{ convex}$$

Example: Ergodic MIMO capacity

$$f(x) = \mathbf{E}_H\{\log\det(HXH^T + I)\} \left(= \int p(H) \log\det(HXH^T + I) dH \right)$$

is concave on $\mathbf{S}_+^n$.

Pointwise maximum:

$$f_1, \ldots, f_m \text{ convex} \Longrightarrow f(x) = \max\{f_1(x), \ldots, f_m(x)\} \text{ convex}$$

Example: Infinity norm

$$f(x) = \|x\|_\infty = \max_{i=1,\ldots,m} |x_i|$$

is convex because $f(x) = \max\{x_1, -x_1, x_2, -x_2, \ldots, x_n, -x_n\}$.

Pointwise supremum:

$$f(x,y) \text{ convex in } x \text{ for each } y \in \mathcal{A} \Longrightarrow \sup_{y \in \mathcal{A}} f(x,y) \text{ convex}$$

Example: Worst-case least squares function

$$f(x) = \max_{E \in \mathcal{E}} \|y - (A+E)x\|_2^2$$

is convex. ($\mathcal{E}$ does not even need to be convex!)

Jensen's Inequality

- The basic inequality for convex f

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

is also called **Jensen's inequality**.

- It can be extended to

$$f\left(\sum_{i=1}^{k} \theta_k x_k\right) \leq \sum_{i=1}^{k} \theta_k f(x_k)$$

where $\theta_1, \ldots, \theta_k \geq 0$, and $\sum_{i=1}^{k} \theta_k = 1$; and to

$$f\left(\int_S p(x) x dx\right) \leq \int_S p(x) f(x) dx$$

where $p(x) \geq 0$ on $S \subseteq \mathbf{dom} f$, and $\int_S p(x)dx = 1$.

Inequalities derived from Jensen's inequality:

- Arithmetic-geometric inequality: $\sqrt{ab} \leq (a+b)/2$ for $a, b \geq 0$
- Hadamard inequality: $\det X \leq \prod_{i=1}^{n} X_{ii}$ for $X \in \mathbf{S}_{+}^{n}$
- Kullback-Leiber divergence: Let $p(x), q(x)$ be PDFs on S,

$$\int_S p(x) \log\left(\frac{p(x)}{q(x)}\right) dx \geq 0$$

- Hölder inequality:

$$x^T y \leq \|x\|_p \|x\|_q$$

  where $1/p + 1/q = 1$, $p > 1$.

Epigraph

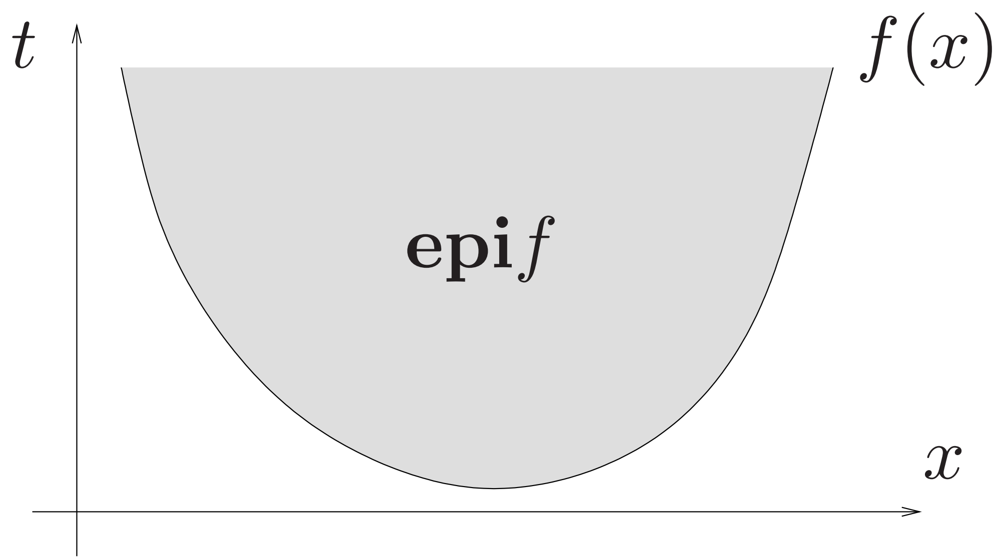
The **epigraph** of f is

$$\mathbf{epi} f = \{(x,t) \mid x \in \mathbf{dom} f, f(x) \le t\}$$

A powerful property:

$$f \text{ convex} \Longleftrightarrow \mathbf{epi} f \text{ convex}$$

e.g., some convexity preserving properties can be proven quite easily by epigraph.

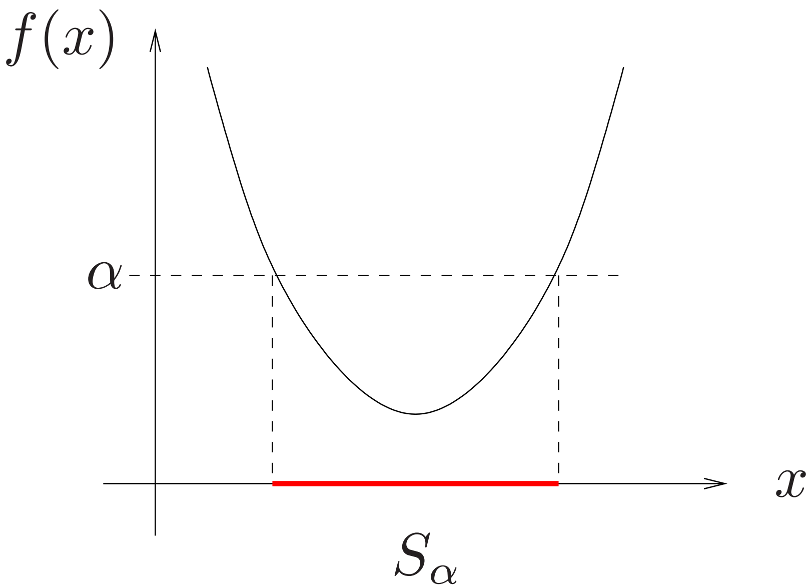


Sublevel Sets

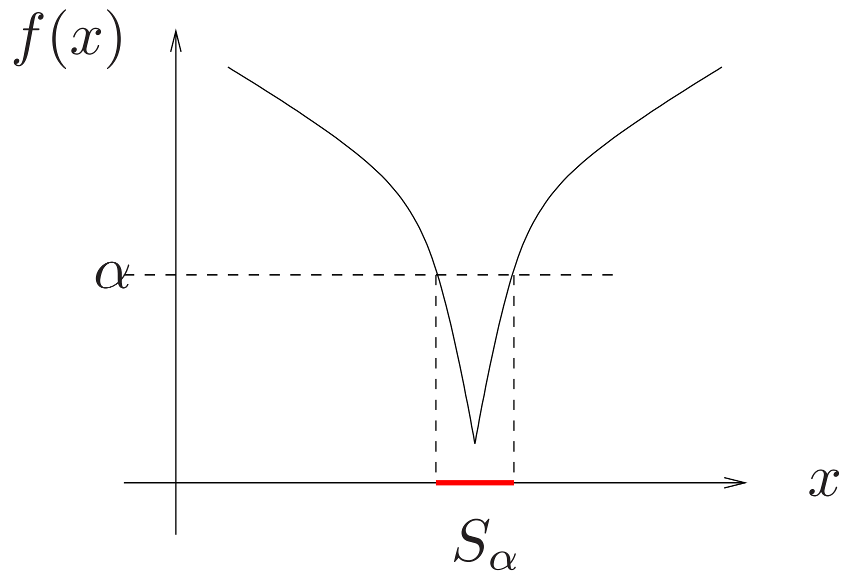
The α-**sublevel set** of f

$$S_\alpha = \{x \in \textbf{dom} f \mid f(x) \leq \alpha\}$$

f convex $\Longrightarrow$ S_α convex for every α, but S_α convex for every $\alpha \not\Longrightarrow f$ convex



convex f and convex S_α



non-convex f but convex S_α

Quasiconvex Functions

$f : \mathbf{R}^n \rightarrow \mathbf{R}$ is called **quasiconvex** (or unimodal) if $\mathbf{dom} f$ is convex and every S_α is convex.

- If f is convex then f is quasiconvex (by the definition).
- f is called **quasiconcave** if $-f$ is quasiconvex.
- f is called **quasilinear** if f is quasiconvex and quasiconcave.

Example: Linear fractional function (useful in modeling SINR)

$$f(x) = \frac{a^T x + b}{c^T x + d}$$

is quasiconvex on $\{x \mid c^T x + d > 0\}$, because

$$S_\alpha = \{x \mid a^T x + b \leq t(c^T x + d)\}$$

is a hyperplane. In fact it is also quasiconcave.